

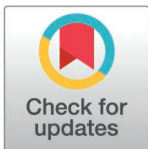
RESEARCH ARTICLE

Exploring heterogeneity in mosquito exposure and attraction and its implications for malaria transmission

Lars Kamber^{1,2}, Aurélien Cavelan^{1,2}, Melissa A. Penny^{1,2,3,4}, Nakul Chitnis^{1,2}, Emma Louise Fairbanks^{1,2,5,6*}

1 Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Allschwil, Switzerland, **2** University of Basel, Basel, Switzerland, **3** The Kids Research Institute Australia, Nedlands, Australia, **4** Centre for Child Health Research, University of Western Australia, Crawley, Australia, **5** The Zeeman Institute for Systems Biology and Infectious Disease Epidemiology Research, Mathematics Institute, University of Warwick, Coventry, United Kingdom, **6** Department of Mathematics, University of Manchester, Manchester, United Kingdom

* emma.fairbanks@manchester.ac.uk



OPEN ACCESS

Citation: Kamber L, Cavelan A, Penny MA, Chitnis N, Fairbanks EL (2026) Exploring heterogeneity in mosquito exposure and attraction and its implications for malaria transmission. *PLoS Comput Biol* 22(9): e1014631. <https://doi.org/10.1371/journal.pcbi.1014631>

Editor: Quirine ten Bosch, Wageningen UR: Wageningen University & Research, NETHERLANDS, KINGDOM OF THE

Received: October 3, 2025

Accepted: July 26, 2026

Published: September 8, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1014631>

Copyright: © 2026 Kamber et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/),

Abstract

Background

Malaria transmission exhibits significant heterogeneity within communities, with small proportions of individuals experiencing disproportionate mosquito exposure.

Methods

This study addresses critical knowledge gaps in characterising and modelling this heterogeneity. Parameterising Bayesian hierarchical models to field data from Burkina Faso, we compared gamma and lognormal distributions for describing heterogeneity in mosquito biting rates. We then implemented this heterogeneity in an individual-based stochastic modelling platform, OpenMalaria, to assess its impact on transmission dynamics.

Findings

The gamma distribution better described the observed field data than the lognormal. This choice has a natural mathematical justification: when individual biting rates follow a gamma distribution and bites occur as a Poisson process, the resulting bite counts follow a negative binomial distribution, which is well-supported empirically for overdispersed count data of this kind. Furthermore, the gamma distribution's lighter tail produces more moderate saturation and immunity effects compared to lognormal-based models, yielding more realistic transmission dynamics. When heterogeneity is introduced to malaria transmission simulations, both prevalence and incidence levels generally decrease across all age groups. Additionally, heterogeneity shifts disease burden towards younger age cohorts and alters the fundamental relationships

which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: Stan files to fit the Bayesian-hierarchical lognormal and gamma models to data are available at <https://github.com/emmairbanks/EntoModels> (DOI: 10.5281/zenodo.21889874). Code for running OpenMalaria simulations is available at <https://github.com/emmairbanks/heterogeneity-mosquito-exposure-OpenMalaria> (DOI: 10.5281/zenodo.21890153). OpenMalaria is an open source software available at: <https://github.com/SwissTPH/openmalaria>.

Funding: All authors were supported by the Bill and Melinda Gates Foundation (INV025569, <https://www.gatesfoundation.org/> to MAP and NC). The funders did not play any role in study design, data analysis, decision to publish or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

between entomological inoculation rate and prevalence/incidence. The integration of appropriate heterogeneity distributions into transmission models substantially improved their ability to reproduce field-observed age-incidence curves.

Interpretation

Our findings highlight the importance of accounting for heterogeneous exposure when modelling malaria transmission, particularly in low-transmission and elimination settings where heterogeneity may be more pronounced.

Author summary

Malaria does not affect everyone equally. Within a community some individuals are bitten by infectious mosquitoes more frequently than others, due to socio-economic factors. This unequal distribution of bites, known as heterogeneity, means that a small proportion of the population often bears a disproportionate share of the transmission burden. This heterogeneity becomes more pronounced as transmission intensity declines, yet it is rarely accounted for in the mathematical models used to guide malaria control policy. In this study, we used statistical methods to analyse field data from Burkina Faso and determine the best mathematical description of how mosquito bites are distributed across individuals. We found that a gamma distribution describes this heterogeneity more accurately than the lognormal distribution. We then incorporated this into OpenMalaria, a widely used malaria transmission model, and showed that accounting for heterogeneity substantially changes predicted patterns of disease burden, particularly across age groups. For several field sites where standard models performed poorly, incorporating heterogeneity improved model fit by up to 95%. As transmission declines and elimination becomes a realistic goal for many countries, ensuring that models accurately capture who is most at risk will be critical for designing interventions that reach those who need them most.

1. Introduction

Malaria prevalence has been observed to vary not only between neighbouring villages, but also within villages [1,2]. Like many diseases, a small proportion of the population often drives malaria transmission within communities. These ‘super-spreaders’ are responsible for a significant portion of the transmission. A study analysing data from 90 African communities revealed that 20% of the population accounts for 80% of infections [3,4]. This heterogeneity arises from multiple factors, including proximity to breeding sites, intervention usage, housing quality, pregnancy status, existing malaria infections, genetic factors and individual human behaviour [5–10]. However, the specific patterns and distributions of exposure heterogeneity at

the household level remain poorly characterised. Understanding these patterns is crucial for accurately modelling transmission dynamics and evaluating control measures [11].

Many studies have observed heterogeneity in mosquito biting rates [12], however, there is a lack of comprehensive quantitative analysis to determine the implications of such heterogeneity. The expected number of infectious mosquito bites a human host receives is often referred to as the entomological inoculation rate (EIR). Data on the EIR is often collected to monitor the effectiveness on interventions, disease burden and vector species compositions [13,14]. Previous research has identified age as a key factor in EIR heterogeneity [15,16]. This is likely influenced in part by body surface area [17], which has been shown to be correlated with biting rates for other vectors and host species [18–20].

Many mathematical models for describing disease transmission dynamics assume homogeneous mixing patterns, which means that all individuals in a population have an equal probability of coming into contact with one another, regardless of their age, location, social connections or behavioural patterns. These assumptions often extend into vector-borne disease models, with each individual equally likely to be bitten by each vector. However, real-world disease transmission patterns exhibit significant heterogeneity, where contact rates and vector biting preferences vary based on factors such as age-structured social mixing, spatial clustering of populations, and vector feeding behaviours [21,22]. Incorporating such heterogeneity into mathematical models can enhance the precision of predictions and policy recommendations.

Whilst the importance of heterogeneous exposure in vector-borne disease transmission was recognised early on, the theoretical implications were first rigorously explored by Dye and Hasibeder [23]. Their work demonstrated that when mosquitoes concentrate on certain hosts, both the basic reproductive number and vectorial capacity may be greater than their values under homogeneous mixing assumptions. Ross and Smith [24] demonstrated that transmission heterogeneity significantly influences all malaria epidemiological outcomes in simulation models, finding that different patterns of heterogeneity produce distinct age-prevalence and incidence curves. Subsequent work has further characterised the implications of heterogeneous exposure across a range of contexts: Woolhouse et al. [3] showed that heterogeneity in transmission has important consequences for the design of control programmes, whilst Cooper et al. [25] demonstrated that Pareto-like rules govern malaria superspreading, with a small proportion of individuals driving a disproportionate share of transmission. The implications of heterogeneous exposure extend to intervention evaluation, with failure to account for heterogeneity shown to bias estimates of vaccine efficacy [11] and to affect predictions of intervention success more broadly [26]. These insights highlight that ignoring heterogeneity in transmission models could lead to underestimating both disease persistence and the control efforts required for elimination [23].

Heterogeneity in biting rates between ages groups is often accounted for in simulations of OpenMalaria, an individual-based stochastic model describing malaria transmission dynamics, reducing the EIR of younger children [27]. However, it is also possible to simulate OpenMalaria assuming an additional heterogeneity in biting rates, where some individuals are more likely to be bitten, independent of age. Here, an OpenMalaria user can add this additional heterogeneity according to an lognormal distribution, specifying a mean and variance.

Our study builds on this foundation by addressing knowledge gaps of heterogeneity in malaria transmission, exploring possible distributions of this heterogeneity and its implications for malaria transmission in different settings. We compare the gamma and lognormal distributions to describe the heterogeneity between the rates at which individuals are bitten. Firstly, we fit previously collected data on DNA fingerprinted mosquito blood meals to compare how well these two distributions describe the data. Then, we conduct OpenMalaria simulations comparing how assumptions of the distribution of EIR heterogeneity affect relationships between different malaria outcomes.

2. Methods

2.1. Ethics statement

The authors confirm that the ethical policies of the journal, as noted on the journal's author guidelines page, have been adhered to. No ethical approval was required.

2.2. Data analysis

Guelbeogo et al. [28] used DNA fingerprinting of human blood meals (matching blood meal DNA to individuals) from wild-caught mosquitoes in Balonghin (health district of Saponé, Burkina Faso) to understand biting heterogeneity between individual human hosts. Here, malaria transmission occurs seasonally between August and December with high prevalence (>80%) during this season [29]. These surveys were performed at the end of the 2013 transmission season (October–December), and at the start (June – July) and peak (September) of the following season in 2014. Each survey indoor mosquito collections were performed in 20–40 households with at least one household member <15 years of age. For each household mosquitoes were collected between 7 and 9 AM by mouth aspiration from walls and ceilings for 5–7 consecutive days within a survey. The head-thoraces of bloodfed mosquitoes were used to identify their species and infection status by PCR.

Here, we use the data from sporozoite positive mosquitoes caught during the study to analyse four datasets; data collected in each of the three surveys and a combined dataset. We only use data from households included in all three surveys, giving a total of 81 individuals. For each individual host we have the number of sporozoite positive blood-fed mosquitoes which were captured and DNA fingerprinted to the individual.

For each dataset we fit two hierarchical Bayesian models to the rate of infectious bites per individual (i); λ_i . We use these Bayesian hierarchical models to consider the variations in the biting rates between individuals. Both models account for variations in biting rates between individuals, but differ in their distributional assumptions: one assumes the individual biting rates follow a lognormal distribution, whilst the other assumes they follow a gamma distribution.

For each model, it is assumed the total number of blood meals found to match each individual follows $Poisson(\lambda_i)$ distribution.

Bayesian inference was performed in Stan [30] in Rstudio [31]. Weakly informed priors were used (Table 1). For each model parameterised, we run four Markov chains with 10,000 iterations, removing the first 5000 for burn-in. The convergence of chains was checked using the diagnostics available within Stan, including the R-statistic and effective sample size.

To compare models, we estimate the marginal log-likelihoods for each model for each dataset, using the *bridgesampling* package [32], averaging over 10 repetitions of the bridge sampling procedure to obtain an empirical estimate of the estimation uncertainty. We calculate posterior model probabilities (PMPs) from the estimated marginal log-likelihoods to quantify the relative support for each model, where the PMP represents the probability of each model being the true data-generating model, assuming equal prior model probabilities.

Lognormal model. For the lognormal model the bite rates ($\lambda_i \in \lambda$) are considered to follow a lognormal distribution. We consider ω which describes how $\log(\lambda)$ varies between individuals, with each individual's $\log(\lambda_i)$ deviating from the

Table 1. Hyperparameters of the lognormal and gamma models. i denotes the individual for parameters that vary between individual hosts.

Parameter	Definition	Prior
Lognormal model		
μ	Population mean rate	$\mathcal{N}(0, 1)$
φ	Population dispersion parameter	Half-Cauchy(0,1)
ω_i	Individual random effects	$\mathcal{N}(0, 1)$
Gamma model		
m	Log of population mean rate	$\mathcal{N}(0, 1)$
$\log(\varphi)$	Log of population dispersion parameter	Half-Cauchy(0,1)
ω_i	Individual random effects	$\mathcal{N}(0, 1)$

<https://doi.org/10.1371/journal.pcbi.1014631.t001>

mean m according to the scale of the standard deviation (σ). We denote the individuals biting rates with a subscript i , for $i = 1, 2, \dots, N$ elements, therefore the individuals rates are given as

$$\lambda_i = \exp(m + \omega_i \sigma).$$

The expected value of the distribution, i.e., the mean of the rates, is therefore

$$\lambda = \mathbb{E}[\lambda_i] = \exp(m + \sigma^2/2). \quad (1)$$

The variance of the rates is

$$\lambda^2(\exp(\sigma^2) - 1) \quad (2)$$

and consequently the standard deviation is

$$\lambda \sqrt{\exp(\sigma^2) - 1}. \quad (3)$$

The coefficient of variation (CV), which provides a standardised measure of dispersion, is given by

$$\sqrt{\exp(\sigma^2) - 1}. \quad (4)$$

Gamma model. For the gamma model, we model the variation in individual biting rates using a Bayesian hierarchical approach with a gamma distribution. The distribution is parameterised by the population mean (μ) and dispersion parameter (φ). The individual rates (λ_i) are derived from a transformation of standard normal random effects (ω_i) to approximate the gamma distribution:

$$\lambda_i = \mu \left(\omega_i \sqrt{\frac{1}{9\varphi}} + \left(1 - \frac{1}{9\varphi} \right) \right)^3$$

This formulation uses the Wilson-Hilferty transformation to convert standard normal random effects into gamma-distributed values. The ω_i terms are individual random effects that follow a standard normal distribution, and the transformation ensures that the resulting distribution of rates approximates a gamma distribution with mean μ and variance μ^2/φ .

This formulation has several desirable properties: (i) the population mean is preserved as μ , (ii) the variation scales with μ (larger means have larger absolute variation) and (iii) the CV is controlled by φ , ensuring that as φ increases, the variation between individuals decreases, matching the gamma distribution's inherent variance structure.

This parameterisation can also be expressed in terms of the traditional gamma shape (α) and rate (β) parameters, where

$$\begin{aligned} \alpha &= \varphi \text{ and} \\ \beta &= \varphi/\mu. \end{aligned}$$

2.3. OpenMalaria simulations

Open Malaria represents a stochastic, individual-based framework for analysing malaria transmission dynamics and evaluating intervention strategies and their public health impact [33–35]. The modelling platform has been applied

to forecast the effectiveness of diverse control strategies, including vector-control and pharmaceutical interventions. OpenMalaria was selected because the epidemiological outcomes of interest (age-structured incidence and prevalence curves) are fundamentally shaped by acquired immunity, which accumulates in an age- and exposure-dependent manner over an individual's lifetime. Simple compartmental models lack the mechanistic representation of immunity necessary to reproduce these dynamics. Within human hosts, OpenMalaria simulates both asexual and sexual parasite stages, immunity acquisition through infection and intervention exposure, drug pharmacokinetics and pharmacodynamics, as well as antimalarial resistance. The model considers clinical disease episodes, both severe and uncomplicated, with individual treatment pathways dictated by a customisable health system architecture. Within OpenMalaria, mosquito dynamics encompass the complete lifecycle and behavioural patterns of malaria-transmitting *Anopheles* species, incorporating key parameters describing host-seeking behaviour, vector survival and responsiveness to insecticidal interventions.

Transmission intensity in OpenMalaria simulations is parameterised through the EIR, the expected number of infectious bites an adult receives annually. In the absence of heterogeneity and interventions, all adults receive the same expected EIR. Introducing heterogeneity creates variation in individual exposure levels, effectively dividing the population into subgroups across a continuous scale of EIRs. Importantly, despite individual variation, the population-level EIR remains equal to the input parameter specified to OpenMalaria. We use the term 'mean EIR' to emphasise this distinction between individual-level and population-level measurements.

To introduce heterogeneity, OpenMalaria assigns each individual an exposure factor sampled from a gamma distribution with mean 1 and a CV which controls the level of heterogeneity. A CV of 0 corresponds to homogeneous transmission. The factor is drawn at model initialisation and birth for each individual. At each time step, the rate which individuals receive bites, $\alpha_{h,i}(t)$, is multiplied by this factor. $\alpha_{h,i}(t)$ itself depends on the age of individuals and their received interventions, such as bed nets, as well as on various characteristics of the mosquito population which are not covered in detail here. The resulting product serves as the rate parameter for a Poisson distribution from which the number of infectious bites the individual receives is randomly sampled each time step. Due to the gamma distribution of the heterogeneity factor, the number of bites for a given age and intervention exposure follow a negative binomial distribution.

When the mean EIR is large, with a high level of heterogeneity, the gamma distribution can lead to some extreme cases of exposure for some individuals. For example, at a CV of 2.5, the 99.99% percentile of the heterogeneity factor is approximately 25. For a mean EIR of 30, this corresponds to at least 710 infectious bites per year for 1 in 10,000 individuals. In order to avoid more extreme values, we truncate the heterogeneity factor at a value of 25 for all simulations. This can result in a change in the mean EIR for CV values larger than 2, with a reduction to 97.5% for a CV of 2.5 and 85% for a CV of 3.5 (S1 Fig). Conversely, high levels of heterogeneity will also result in very low heterogeneity factors for some individuals, effectively excluding them from malaria transmission entirely. To assess the proportion of the population at risk, we track the age of each individual's first malaria infection. If individuals are not infected before the age of 60, we consider them to be excluded from malaria transmission; this threshold was chosen as a conservative upper bound representing a typical lifetime of exposure opportunity within the simulated population, rather than as a biological claim about the susceptibility of older individuals.

We consider the EIR-prevalence and EIR-incidence relationships alongside age-prevalence and age-incidence curves to assess the impact of heterogeneity on these fundamental epidemiological relationships at equilibrium. To ensure equilibrium conditions, each OpenMalaria simulation includes a 100-year burn-in period, allowing the system to reach a stable state where population immunity fully reflects long-term exposure patterns. This burn-in guarantees that all reported results and figures represent true steady-state values, eliminating any transient dynamics that might otherwise confound the analysis. We explore transmission intensities ranging from mean EIRs of 1–300, combined with heterogeneity levels parameterised through CVs from 0 to 2.5. For simplicity, we assume no seasonality in malaria transmission. The age distribution of the human population is taken from the parameterisation used to fit the base parameters of OpenMalaria [36]

and effective coverage for uncomplicated malaria cases is 37% [37]. A detection threshold of 40 parasites/ L of blood is used for prevalence calculations. The mosquito species is *An. funestus* with a blood index of 0.98.

In a second analysis, we explore how introducing heterogeneity can improve the fit of OpenMalaria to age-incidence and age-prevalence curves derived from field data [38]. The considered curves are used for fitting OpenMalaria parameters [39], but a good fit cannot be achieved for some of the locations under the assumption of homogenous biting exposure. For these locations, there was likely a high degree of heterogeneity, but the extent of heterogeneity is unknown and a method for introducing heterogeneity into OpenMalaria has not been previously established, which is why heterogeneity has not been included in the fitting process. The parameterisation of mosquito bionomics, population structure, interventions, seasonality and many other factors for the scenarios are adapted to each site of data collection [39]. We run all the scenarios used in the fitting process of these curves including those where we do not necessarily suspect heterogeneity to be present. We vary the CV parameter to determine the level of heterogeneity that minimises the sum of squared distances from model output to the data. The optimal CV value is determined independently for the age-prevalence and age-incidence curves for each scenario. This allows us to assess the agreement of the estimated heterogeneity between the two curves for a given location. To quantify the magnitude of change introduced by heterogeneity, we calculate the sum of squared differences between model outputs under homogeneous transmission and those with the best-fit CV values.

3. Results

3.1. Data analysis

The marginal log-likelihood estimates consistently favoured the gamma model across all temporal subsets of the data, as well as the combined dataset (Table 2). The estimated values of the marginal log-likelihoods varied substantially across the temporal subsets, with the peak period showing markedly higher values (76.679 and 83.766 for lognormal and gamma models, respectively) compared to the start and end periods. This pattern suggests better model fit during the peak period.

The precision of the marginal log-likelihood estimates, as measured by the interquartile range, ranged from 0.011 to 0.026. There was no consistent pattern in which the model showed a lower estimation uncertainty.

The PMPs provide a direct measure of the relative support for each model in the data. The PMPs showed a consistent pattern, with the gamma model achieving probabilities between 0.994 and 1 in all datasets, indicating stronger support for this model specification.

Fig 1 shows that the fitted gamma distribution describes the data better than the lognormal. The composite estimates for the mean and CV of the distributions are much more uncertain for the lognormal distribution (S2 Fig). The posterior

Table 2. The median and interquartile range width of the marginal log-likelihood estimates for the lognormal and gamma models, as well as the mean posterior model probabilities (PMPs). Results are given to 3 Decimal Places.

Dataset	Model	Median	Interquartile range width	PMPs
Start	Lognormal	40.973	0.011	0.000
	Gamma	49.917	0.018	1.000
Peak	Lognormal	76.679	0.014	0.001
	Gamma	83.766	0.026	0.999
End	Lognormal	42.732	0.017	0.003
	Gamma	48.512	0.015	0.997
Combined	Lognormal	113.501	0.022	0.006
	Gamma	118.594	0.017	0.994

<https://doi.org/10.1371/journal.pcbi.1014631.t002>

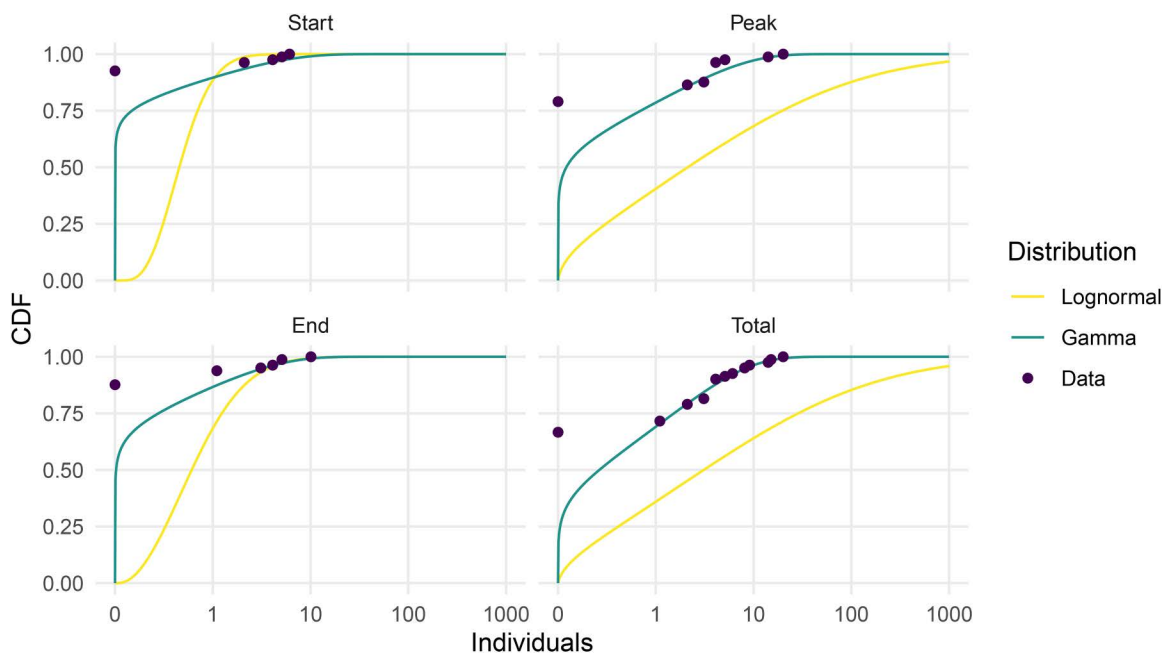


Fig 1. The cumulative density function of mosquito exposure over the human individuals of the fitted lognormal and gamma models and the empirical data.

<https://doi.org/10.1371/journal.pcbi.1014631.g001>

median for CV ranged from 5.44–12.18 and 1.72–2.14 under the lognormal and gamma distribution assumptions, respectively. For the gamma distribution the CV is lower for the start and end of the seasons, compared to the peak.

The lognormal distribution notably underestimates the proportion of individuals with low exposure values, particularly during peak transmission and when considering the combined dataset. This is because the lognormal distribution's heavier tail forces a slower rise in the CDF to accommodate even a few individuals with higher exposure, sacrificing fit quality for the majority of the population. This sensitivity to outliers is clearly demonstrated when comparing the 'Start' and 'End' periods: despite similar exposure patterns for most individuals (fewer than 5 bites), the presence of a single individual receiving 10 bites during the 'End' period causes a noticeable downward shift in the lognormal CDF. In contrast, the gamma distribution maintains a better balanced fit across the entire range of the data even when accommodating occasional outliers. This robustness to extreme values whilst preserving accuracy for the majority of observations confirms our model selection results that consistently favoured the gamma distribution across all temporal periods.

3.2. Effects of heterogeneity at population level

In the absence of heterogeneity, OpenMalaria estimates for malaria prevalence and incidence at the population level reach saturation at relatively low transmission intensities (Figs 2A and 3A). Parasite prevalence peaks at 59% with a mean EIR of 50, with the 50% threshold crossed at a substantially smaller mean EIR of 10. Clinical incidence follows a different pattern, reaching its maximum of 1.25 cases per person per year at an EIR of 9, before declining as the mean EIR increases further. This post-peak reduction occurs because increased exposure strengthens population immunity, thereby reducing clinical episodes. Importantly, this declining pattern is not observed for prevalence, as immunity, whilst reducing parasitaemia, does not eliminate it completely.

When introducing heterogeneity with the gamma distribution, more individuals experience an EIR below the mean than above it due to the distribution's right-skewed nature. As an example, we compare prevalence across all ages between

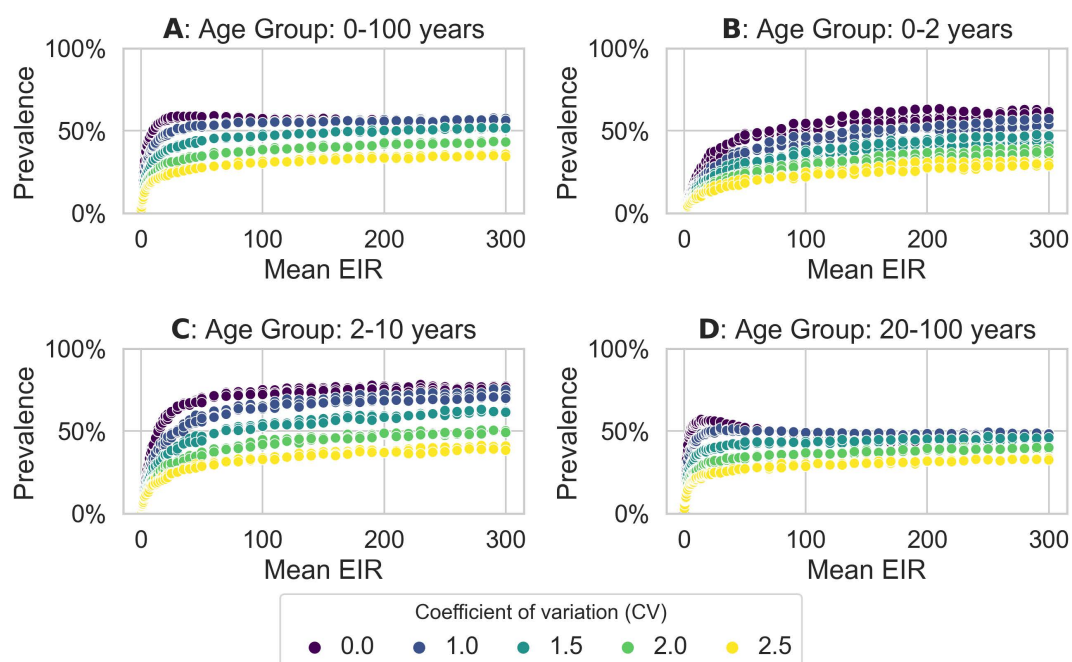


Fig 2. Parasite prevalence across the entire population (A) and specific age groups (B-D) at varying transmission intensities (mean EIRs). Each point represents a simulation. The CV=0 curves can be interpreted as approximations of the EIR-prevalence relationships at the individual level.

<https://doi.org/10.1371/journal.pcbi.1014631.g002>

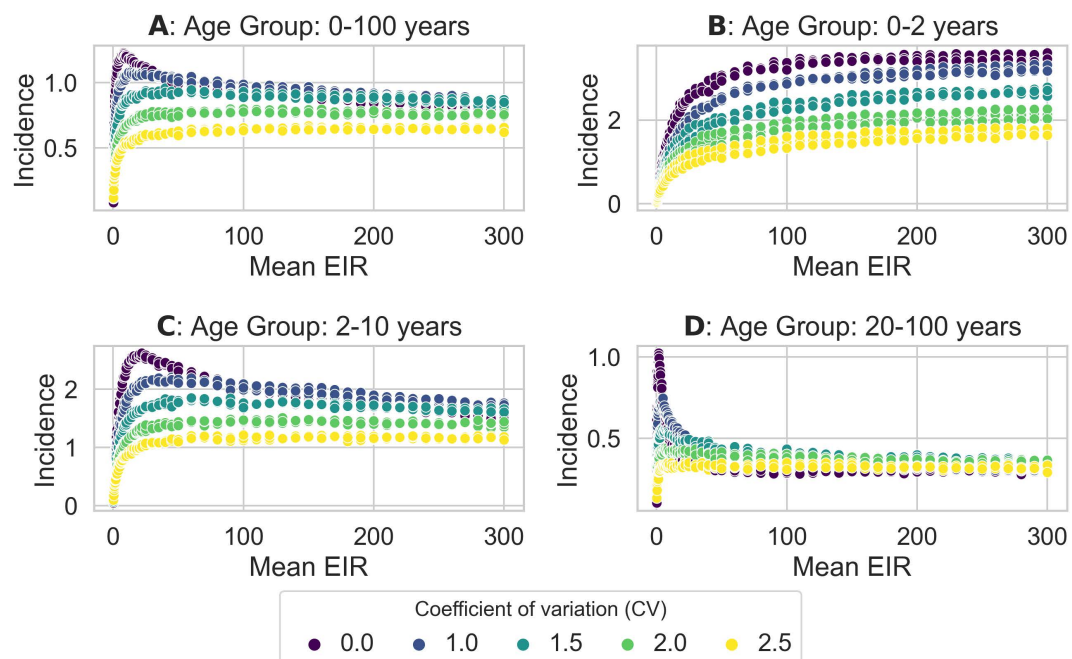


Fig 3. Clinical incidence across the entire population (A) and specific age groups (B-D) at varying transmission intensities (mean EIRs). Each point represents a simulation. The CV=0 curves can be interpreted as approximations of the EIR-prevalence relationships at the individual level.

<https://doi.org/10.1371/journal.pcbi.1014631.g003>

a homogeneous scenario ($CV=0$) and a heterogeneous scenario ($CV=1$) at a mean EIR of 100 (S3 Fig). As the majority of individuals are shifted below the mean EIR of 100, they experience individual EIRs that result in similar or lower prevalence levels. Whereas, individuals shifted above the mean EIR of 100 experience only marginally higher prevalence. Overall, since there are more individuals with an EIR lower than the mean, this yields lower overall prevalence compared to homogeneous exposure. This effect holds for any mean EIR and increases with heterogeneity, as more individuals experience low EIRs with an increase in CV. This results in prevalence remaining below 40% even at mean EIRs as high as 300 under highest studied heterogeneity conditions (Fig 2).

At the population level, more heterogeneity also results in lower clinical incidence, with a notable exception at low levels of heterogeneity. As an example, considering a mean EIR of 100 and $CV=1$ (S4 Fig), the redistribution of individuals below the mean EIR drives them towards higher incidence levels due to reduced immunity development. Through this mechanism, low heterogeneity produces slightly higher overall incidence values in high transmission settings (Fig 3A). However, with higher heterogeneity levels, enough individuals are driven towards sufficiently low EIRs that the effect of reduced exposure outweighs the impact of reduced immunity.

3.3. Effects of heterogeneity across age groups

The EIR-prevalence and EIR-incidence curves show very different relationships across age groups, both in homogeneous and heterogeneous transmission scenarios. In the absence of heterogeneity, adults exhibit the most distinctive pattern (Fig 3D), with clinical incidence sharply peaking at a low mean EIR of 1.75, before rapidly declining as the mean EIR increases. This pattern reflects immunity dynamics; higher transmission intensity leads to greater accumulated immunity in adults, subsequently reducing clinical episodes. In contrast, younger age groups display increasingly monotonic relationships (Fig 3B and 3C), where higher mean EIRs consistently correspond to increased clinical incidence. Prevalence exhibits similar age-dependent patterns, with younger groups showing trends toward monotonicity (Fig 2B, 2C, and 2D). As the level of heterogeneity is increased, both prevalence and incidence relationships increasingly shift toward monotonicity regardless of age.

As demonstrated in Section 3.2, both overall incidence and prevalence generally decrease with increasing heterogeneity across all mean EIR values. Fig 4A shows that the higher overall incidence observed in settings with lower heterogeneity accumulates primarily in younger age groups, while incidence rates among older individuals remain similar regardless of heterogeneity level. Under homogeneous transmission, most individuals experience their first infections at a young age, acquiring immunity accordingly. As mean EIR increases, these infections occur progressively earlier in life, causing age-incidence curves to decline more steeply with age. For prevalence, lower heterogeneity consistently produces higher estimates across most age groups, with an exception occurring in older individuals exposed to high transmission intensity under low heterogeneity conditions (Fig 4B). This contrasting pattern between prevalence and incidence can be attributed to acquired immunity, which prevents clinical episodes but not infections.

When comparing the age-incidence curves at a given prevalence instead of a given mean EIR, disease burden shifts towards younger age groups as heterogeneity increases (Fig 4C). This pattern emerges because achieving equivalent population-level prevalence requires significantly higher mean EIRs under greater heterogeneity. For example, an overall prevalence of 35-37.5% requires a mean EIR of 4 at $CV=0$, increasing to a mean EIR of 7 at $CV=1$, and rising substantially to mean EIRs of 69 and 275 at $CV=2$ and $CV=2.5$, respectively. Under homogeneous transmission ($CV=0$), a relatively low mean EIR results in infrequent infections across the entire population, producing gradual immunity development and allowing infections to occur across a broader age range. In contrast, at high heterogeneity ($CV=2.5$), the high mean EIR concentrates exposure in a subpopulation that rapidly accumulates infections at an early age, quickly developing immunity and experiencing fewer infections as they age. This pattern — shifting incidence towards younger ages with increasing heterogeneity — persists even when accounting for the effective exclusion of part of the population from infection risk under high heterogeneity conditions (S5, S6 and S7 Figs).

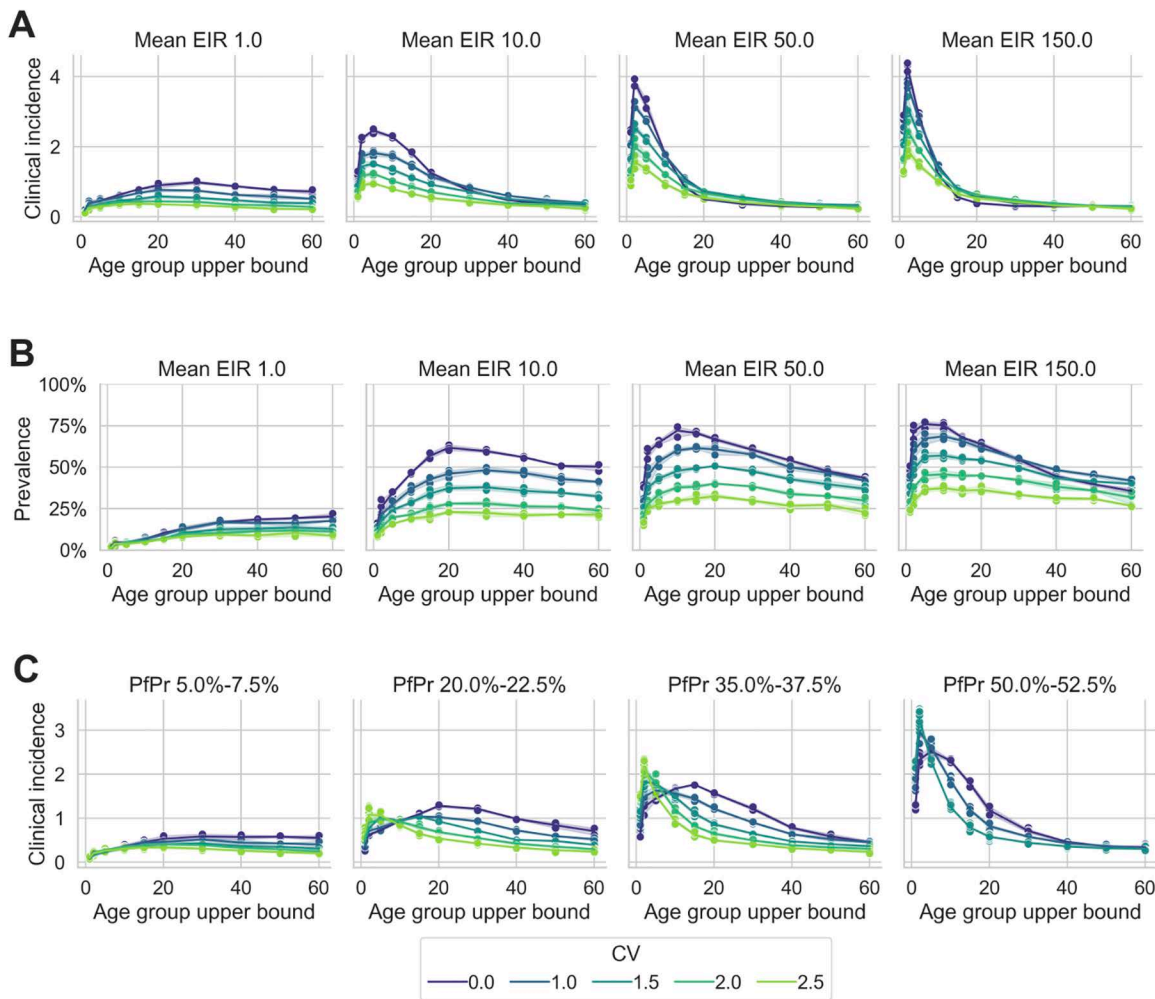


Fig 4. Age-incidence (A) and age-prevalence curves (B) at varying transmission intensities (mean EIRs), and age-incidence curves for varying levels of prevalence (C).

<https://doi.org/10.1371/journal.pcbi.1014631.g004>

For high levels of heterogeneity ($CV \geq 2$), the population at risk is below 50% for low mean EIRs, gradually increasing to approximately 80% at higher mean EIRs, reaching saturation. In contrast, for lower levels of heterogeneity, the entire population is at risk of infection. This leads to an increase in incidence and prevalence for any given mean EIR when adjusting for the proportion of the population at risk. The age-incidence curves corrected for this effect exhibit the same pattern observed in uncorrected analyses (S5, S6 and S7 Figs): higher heterogeneity consistently shifts disease burden towards younger age groups. This demonstrates that the observed pattern is not just a consequence of restricting transmission to a subpopulation, but represents an effect of heterogeneity *within* the malaria-affected subpopulation. In contrast to the uncorrected version, prevalence levels of 65% and higher can be achieved when considering only the population at risk.

3.4. Parametrising heterogeneity from age-prevalence and age-incidence curves

For several tested scenarios, incorporating heterogeneity substantially improves the alignment between simulated age-incidence curves and field observations. Fig 5 shows the three scenarios where including heterogeneity the yielded the most improvements. For these scenarios, the sum of squared distances — measuring deviation between observed

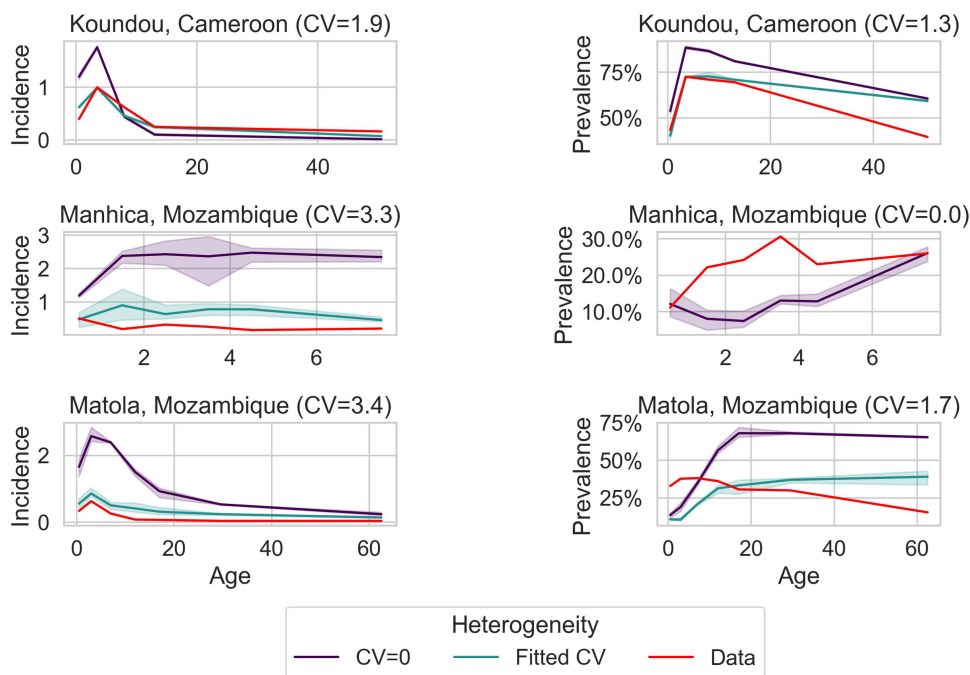


Fig 5. Age-incidence (left column) and age-prevalence curve (right column) of three scenarios that were most improved by added heterogeneity. Fitted CV curves show predictions with the model parameterised using the best-fit CV value, which is also given in the brackets.

<https://doi.org/10.1371/journal.pcbi.1014631.g005>

data and model predictions — decreased by approximately 95%. The optimal heterogeneity levels varied considerably: the scenarios from Manhica and Matola (both Mozambique) required very high heterogeneity, while the scenario from Koundou, Cameroon showed best results with more moderate heterogeneity. For prevalence, the best-fit CV values was substantially smaller with a much more moderate improvement in the fit, with a CV of 0 producing the best fit for scenario Manhica. For all scenarios tested, [S8](#) and [S9 Figs](#) show model outputs when the model is parameterised using the CV value which produced the best fit and [S1 Table](#) details the improvement by introducing heterogeneity.

4. Discussion

This study advances our understanding of heterogeneity in malaria transmission beyond the foundational theoretical work of Dye and Hasibeder [23]. Whilst they established that heterogeneous biting increases the basic reproductive rate and demonstrated this effect using field data, our work provides a rigorous statistical framework for determining which probability distributions best characterize this heterogeneity. Integrating heterogeneous biting exposure into OpenMalaria leads to substantial changes in basic equilibrium epidemiological relationships.

Within OpenMalaria, the effects of heterogeneity are driven by two interrelated mechanisms: saturation effects and immunity development. Saturation effects arise when individuals receive multiple infectious bites. Beyond a certain threshold, additional bites contribute little to further disease incidence. If an individual is already experiencing a clinical malaria episode, subsequent infections progressing to the blood stage are not counted as separate episodes. Thus, whilst the risk of developing an infection increases with the number of infectious bites, each additional bite has progressively less impact on disease risk. This saturation effect, arising from the concave relationship between EIR and infection risk, was previously demonstrated theoretically [40] and empirically across African field sites [4], albeit without the immunity dynamics modelled here.

Immunity effects add further complexity to transmission dynamics. At equilibrium, where individuals receive consistent average exposure over extended periods, those exposed to higher bite rates develop stronger immunity. Consequently, increasing bite exposure for an individual typically results in fewer clinical episodes over time, particularly in adults where immunity has accumulated over years. This creates an inverse relationship between accumulated exposure and infection risk.

The interaction between saturation and immunity explains the observed effects of heterogeneity in OpenMalaria simulations. Generally, higher heterogeneity reduces population-level incidence due to saturation effects amongst heavily exposed individuals. However, immunity can reverse this effect in adults for certain levels of heterogeneity. Under homogeneous exposure, most adults develop sufficient immunity to substantially reduce disease incidence. With heterogeneity, whilst individuals with high exposure also develop strong immunity, some adults remain relatively unexposed and fail to develop adequate immunity for disease prevention. These individuals experience occasional infections throughout adult life, thereby increasing overall adult incidence despite the general trend towards lower population-level transmission.

We explored the effect of heterogeneity over a wide range of transmission settings, including those with very high EIR. Previous research has shown greater heterogeneity in lower EIR settings [16,25], with pockets of infections. Whilst including the combination of high transmission activity and high heterogeneity in the analysis provides a comprehensive assessment of the effects of heterogeneity in OpenMalaria, such settings are less likely to occur in natural contexts. We can implement decreasing heterogeneity with increasing EIR in OpenMalaria, but currently have insufficient data to parameterise this relationship appropriately.

Seasonality was excluded from the primary simulations in order to isolate the effects of individual-level heterogeneity and avoid confounding interactions between the two. In the presence of seasonal forcing, the system does not reach a static equilibrium but rather a stable periodic solution, returning to the same annual cycle each year. Our 100-year burn-in period is sufficient to ensure convergence to this periodic steady state. The interaction between seasonality and heterogeneity is complex: immunity acquisition depends on the timing and intensity of seasonal transmission peaks, and our analysis showed that the inverse correlation between EIR and heterogeneity may not only be applicable regionally, but also through seasonal variations in transmission. A larger CV was observed during the start and end of the season, when transmission levels were lower. This is consistent with previous modelling work demonstrating that heterogeneity in exposure plays a particularly important role in low-transmission settings, where it can substantially affect the success of interventions [26]. A full characterisation of the joint effects of seasonality and individual-level heterogeneity on transmission dynamics therefore represents an important direction for future work.

By comparing gamma and lognormal distributions using Bayesian hierarchical models and modern model selection techniques, we show that the gamma distribution better captures the observed patterns of exposure heterogeneity. Whilst the lognormal distribution has theoretical appeal for modelling extreme values through its heavy tail, it systematically overestimates the proportion of individuals receiving very high exposure whilst underestimating those with low exposure for our datasets. Although the gamma distribution lacks the heavy tail of the lognormal, it provides a more accurate representation of the actual distribution of biting exposure in the population.

The absence of heavy tails in the gamma distribution has important consequences for modelling saturation and immunity effects. Without the extreme outliers characteristic of heavy-tailed distributions, fewer individuals reach the saturation threshold where additional bites yield diminishing returns on disease incidence. Simultaneously, the gamma distribution's more moderate extreme values mean that highly exposed individuals accumulate immunity more gradually, reducing the high contrast between high and low exposure groups. This modulation of both saturation and immunity effects likely produces more realistic transmission dynamics compared to the polarised populations generated by heavy-tailed distributions. However, even after right-truncation, the gamma distribution can result in some extreme differences in exposure between individuals under high heterogeneity. Possible approaches to this problem are additional left-truncation setting a minimal exposure or alternative distributions that lead to less extreme values, for example a triangular distribution or a nonparametric leading to a 80/20 distribution of biting exposure.

Many studies that estimate EIRs report averages and confidence intervals over the duration of the study. More detailed data for an individual level would help further our understanding of these distributions. To understand heterogeneity at the individual level, studies that perform human landing catches or DNA fingerprinted of resting, blood-fed mosquitoes would be the most informative. While DNA fingerprinting of mosquito blood meals provides valuable insights into biting heterogeneity, potential sampling biases warrant consideration. Our analytical approach assumes that the absence of matched blood meals for certain individuals genuinely reflects lower biting rates. However, methodological limitations could artificially inflate heterogeneity estimates if mosquitoes that fed on particular individuals are systematically underrepresented in collections. This might occur due to spatial variation in mosquito resting behaviours, household-specific factors affecting mosquito collection efficiency, or stochastic sampling effects in households with few mosquitoes.

In this study we demonstrated that individual-level heterogeneity can be directly parameterised from data. We also showed that heterogeneity can allow the model to fit age-incidence curves from field data that could not be reproduced under homogeneous exposure. Whilst heterogeneity was documented during data collection from studies in Manhica, Mozambique and Dielmo, Senegal (scenarios s30 and s32 respectively in [S1 Table](#)), this was not the case for all scenarios showing improved fit. Importantly, heterogeneity primarily improves the fit by lowering simulated incidence levels, though such reductions could also arise from factors such as incomplete case reporting. Fitting to age-prevalence curves yielded markedly different results: heterogeneity estimates were substantially lower and improvements were minimal. Additionally, there was no statistically significant correlation between the heterogeneity fitted to age-prevalence and age-incidence curves ([S1 Table](#)). This could be due to two possible reasons. First, this inconsistency might suggest that heterogeneity is not the primary driver of the observed age-specific patterns in both indicators. Alternatively, the discrepancy could reflect fundamental differences in how these indicators are measured: prevalence data is typically obtained from cross-sectional surveys, whilst incidence relies on active surveillance systems. These measurement approaches are subject to different biases, possibly skewing results in opposing directions.

The higher heterogeneity observed during periods of lower transmission suggests that control interventions may face greater challenges in elimination settings than previously recognised. The incorporation of appropriate heterogeneity distributions in transmission models enhances their appropriateness for modelling transmission as we approach elimination settings. As transmission intensity declines, exposure becomes increasingly concentrated among a smaller proportion of the population [3,25]. This suggests that standard interventions distributed uniformly across a community may be less efficient in elimination settings, where high-exposure individuals drive a disproportionate share of transmission [11,25]. Future work should explore methods for identifying and targeting these individuals to maximise the impact of limited resources.

The incorporation of appropriate heterogeneity distributions into transmission models such as OpenMalaria enhances their suitability for guiding policy in low-transmission and elimination settings. Our results suggest that models assuming homogeneous exposure may systematically misrepresent the age distribution of disease burden and the relationship between EIR and epidemiological outcomes in these settings. Countries currently using OpenMalaria to inform elimination strategies should consider incorporating heterogeneity, particularly where observed age-incidence curves deviate from model predictions under homogeneous assumptions, as demonstrated here for sites in Mozambique and Cameroon.

Under heterogeneous transmission, a proportion of highly exposed individuals accumulate infections and develop immunity earlier in life, shifting the peak of age-incidence curves towards younger ages when comparing settings of equivalent prevalence. Conversely, less exposed individuals may fail to develop adequate immunity, potentially increasing clinical incidence in older age groups. Models that do not account for this heterogeneity may therefore misrepresent the age distribution of disease burden, with implications for the targeting of age-specific interventions such as seasonal malaria chemoprevention. Future work should explore how incorporating individual-level heterogeneity affects predictions of intervention impact in transmission models.

Hasibeder and Dye [41] examined how spatial clustering of vectors and hosts could further amplify transmission beyond individual-level heterogeneity, showing that strong associations between particular groups of mosquitoes and hosts could alter transmission dynamics. An extension of the present work would be the explicit inclusion of spatial heterogeneity in OpenMalaria through distinct mosquito populations. Similarly, the temporal variation in heterogeneity observed across transmission seasons, with greater heterogeneity during periods of lower transmission, suggests that static parameterisations of heterogeneity may be insufficient. Seasonal models that allow heterogeneity to vary with transmission intensity could provide more accurate representations of exposure dynamics, particularly in highly seasonal settings approaching elimination.

In summary, this study demonstrates that heterogeneous mosquito exposure is not merely a theoretical consideration but has substantive consequences for how malaria transmission is modelled and how control strategies are designed. The gamma distribution provides a statistically rigorous and practically meaningful characterisation of individual-level exposure heterogeneity, and its incorporation into OpenMalaria substantially improves model fit to field observations. As the malaria community works towards elimination, the concentration of transmission among a small proportion of highly exposed individuals means that models assuming homogeneous exposure risk misrepresenting both the epidemiological burden and the impact of interventions. Incorporating heterogeneity into transmission models such as OpenMalaria should therefore be considered a priority, particularly in low-transmission settings where its effects are most pronounced and the consequences of misspecification most consequential for policy.

Supporting information

S1 Fig. Effect on the mean of truncating a gamma distribution with original mean 1 at a truncation value of 25.

Up to a coefficient of variation (CV) of 2, the effect on the mean is limited. At higher CVs, the mean can be substantially altered due to truncation. The values were calculated once using numerical integration of the gamma probability density function and once through random number draws to ensure robustness of the results.

(TIFF)

S2 Fig. Parameter estimate median (dot) and interquartile range (line) for each model and dataset.

(TIFF)

S3 Fig. Prevalence of clinical malaria cases across the entire population (A) and specific age groups (B-D) at varying transmission intensities (mean EIRs). Each point represents a simulation. The blue line represents the distribution of EIRs of individuals at a mean EIR of 100 (black dotted line) and a CV of 1.0. The CV=0 curves can be interpreted as approximations of the EIR-prevalence relationships at the individual level.

(TIFF)

S4 Fig. Clinical malaria incidence across the entire population (A) and specific age groups (B-D) at varying transmission intensities (mean EIRs). Each point represents a simulation. The blue line represents the distribution of EIRs of individuals at a mean EIR of 100 (black dotted line) and a CV of 1.0. The CV=0 curves can be interpreted as approximations of the EIR-prevalence relationships at the individual level.

(TIFF)

S5 Fig. Proportion of individuals which have experienced their first infection at each age. This fraction can be non-monotonous due to stochastic migration effects contained in OpenMalaria to keep the age distribution constant.

(TIFF)

S6 Fig. Proportion of individuals that experienced at least one infection at age 60, used as a proxy for the proportion of the population at risk. This corresponds to the last data point for each EIR as plotted in [S5 Fig.](#)

(TIFF)

S7 Fig. Age-incidence curve over prevalences corrected for the fact that under high levels of heterogeneity, some individuals do not experience infections over their lifetimes. The incidence and the prevalence from the overall population were divided by the proportion of the population at risk.

(TIFF)

S8 Fig. Age-incidence curves collected by Battle et al. used for the fitting of OpenMalaria. Heterogeneity was fitted through the coefficient of variation (CV) such that the distance between data and the curve generated by OpenMalaria was minimised as measured by sum of squares. The fitted CV value is given in brackets behind the scenario names. The improvement of the fit to the data with heterogeneity compared to no heterogeneity is given in [S1 Table](#).

(TIFF)

S9 Fig. Age-prevalence curves collected by Battle et al. used for the fitting of OpenMalaria. Heterogeneity was fitted through the coefficient of variation (CV) such that the distance between data and the curve generated by OpenMalaria was minimised as measured by sum of squares. The fitted CV value is given in brackets behind the scenario names. The improvement of the fit to the data with heterogeneity compared to no heterogeneity is given in [S1 Table](#).

(TIFF)

S1 Table. Fitted heterogeneity to age-incidence curves (middle section of the table) and age-prevalence curves (right section of the table). The improvement columns show the reduction in the sum squared error between the curves without heterogeneity and the curves with fitted heterogeneity. If the fitted CV equals 0 (i.e., no heterogeneity), the improvement is 0%. The fitted CV for incidence and prevalence has a weak, nonsignificant correlation of 0.18. Spearman rank correlation is also nonsignificant.

(TIFF)

Acknowledgments

Calculations were performed at sciCORE (<http://scicore.unibas.ch/>) scientific computing center at University of Basel. ELF acknowledges support from a University of Manchester Healthier Futures Fellowship and L'Oréal-UNESCO For Women in Science Award.

Author contributions

Conceptualization: Lars Kamber, Aurélien Cavelan, Melissa A. Penny, Nakul Chitnis, Emma Louise Fairbanks.

Data curation: Emma Louise Fairbanks.

Formal analysis: Lars Kamber, Emma Louise Fairbanks.

Funding acquisition: Melissa A. Penny, Nakul Chitnis, Emma Louise Fairbanks.

Investigation: Lars Kamber, Emma Louise Fairbanks.

Methodology: Lars Kamber, Aurélien Cavelan, Nakul Chitnis, Emma Louise Fairbanks.

Project administration: Emma Louise Fairbanks.

Resources: Melissa A. Penny, Nakul Chitnis.

Software: Lars Kamber, Aurélien Cavelan, Emma Louise Fairbanks.

Supervision: Nakul Chitnis.

Validation: Lars Kamber, Emma Louise Fairbanks.

Visualization: Lars Kamber, Emma Louise Fairbanks.

Writing – original draft: Lars Kamber, Emma Louise Fairbanks.

Writing – review & editing: Lars Kamber, Nakul Chitnis, Emma Louise Fairbanks.

References

- Greenwood BM. The microepidemiology of malaria and its importance to malaria control. *Trans R Soc Trop Med Hyg.* 1989;83 Suppl:25–9. [https://doi.org/10.1016/0035-9203\(89\)90599-3](https://doi.org/10.1016/0035-9203(89)90599-3) PMID: 2576161
- Rumisha SF, Smith T, Abdulla S, Masanja H, Vounatsou P. Modelling heterogeneity in malaria transmission using large sparse spatio-temporal entomological data. *Glob Health Action.* 2014;7:22682. <https://doi.org/10.3402/gha.v7.22682> PMID: 24964782
- Woolhouse EJ, Dye C, Etard J-F, Smith T, Charlwood JD, Garnett GP, et al. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proc Natl Acad Sci USA.* 1997;94(1):338–42. <https://doi.org/10.1073/pnas.94.1.338>
- Smith DL, Dushoff J, Snow RW, Hay SI. The entomological inoculation rate and *Plasmodium falciparum* infection in African children. *Nature.* 2005;438(7067):492–5. <https://doi.org/10.1038/nature04024> PMID: 16306991
- Mackinnon MJ, Mwangi TW, Snow RW, Marsh K, Williams TN. Heritability of malaria in Africa. *PLoS Med.* 2005;2(12):e340. <https://doi.org/10.1371/journal.pmed.0020340>
- Bannister-Tyrrell M, Verdonck K, Hausmann-Muela S, Gryseels C, Muela Ribera J, Peeters Grietens K. Defining micro-epidemiology for malaria elimination: systematic review and meta-analysis. *Malar J.* 2017;16(1):164. <https://doi.org/10.1186/s12936-017-1792-1> PMID: 28427389
- Ellwanger JH, Cardoso JDC, Chies JAB. Variability in human attractiveness to mosquitoes. *Curr Res Parasitol Vector Borne Dis.* 2021;1:100058. <https://doi.org/10.1016/j.crvbd.2021.100058> PMID: 35284885
- Mwema T, Lukubwe O, Joseph R, Maliti D, litula I, Katokele S, et al. Human and vector behaviors determine exposure to *Anopheles* in Namibia. *Parasit Vectors.* 2022;15(1):436. <https://doi.org/10.1186/s13071-022-05563-6>
- Buckee C, Noor A, Sattenspiel L. Thinking clearly about social aspects of infectious disease transmission. *Nature.* 2021;595(7866):205–13. <https://doi.org/10.1038/s41586-021-03694-x> PMID: 34194045
- Aidoo EK, Aboagye FT, Botchway FA, Osei-Adjei G, Appiah M, Duku-Takyi R, et al. Reactive case detection strategy for malaria control and elimination: A 12 year systematic review and meta-analysis from 25 malaria-endemic countries. *Trop Med Infect Dis.* 2023;8(3):180. <https://doi.org/10.3390/tropicalmed8030180>
- White MT, Griffin JT, Drakeley CJ, Ghani AC. Heterogeneity in malaria exposure and vaccine response: implications for the interpretation of vaccine efficacy trials. *Malar J.* 2010;9:82. <https://doi.org/10.1186/1475-2875-9-82> PMID: 20331863
- Mbogo CM, Mwangangi JM, Nzovu J, Gu W, Yan G, Gunter JT, et al. Spatial and temporal heterogeneity of *Anopheles* mosquitoes and *Plasmodium falciparum* transmission along the Kenyan coast. *Am J Trop Med and Hyg.* 2003;68(6):734–42.
- Beier JC, Killeen GF, Githure JI. Short report: entomologic inoculation rates and *Plasmodium falciparum* malaria prevalence in Africa. *Am J Trop Med Hyg.* 1999;61(1):109–13. <https://doi.org/10.4269/ajtmh.1999.61.109> PMID: 10432066
- Shaukat AM, Breman JG, McKenzie FE. Using the entomological inoculation rate to assess the impact of vector control on malaria parasite transmission and elimination. *Malar J.* 2010;9:122. <https://doi.org/10.1186/1475-2875-9-122> PMID: 20459850
- Molineaux L, Gramiccia G, World Health Organization. The Garki project: research on the epidemiology and control of malaria in the Sudan savanna of West Africa. World Health Organization; 1980.
- Mugenyi L, Abrams S, Hens N. Estimating age-time-dependent malaria force of infection accounting for unobserved heterogeneity. *Epidemiol Infect.* 2017;145(12):2545–62. <https://doi.org/10.1017/S0950268817001297> PMID: 28677517
- Port GR, Boreham PFL, Bryan JH. The relationship of host size to feeding by mosquitoes of the *Anopheles gambiae* giles complex (Diptera: Culicidae). *Bull Entomol Res.* 1980;70(1):133–44. <https://doi.org/10.1017/s0007485300009834>
- Viennet E, Garros C, Gardès L, Rakotoarivony I, Allène X, Lancelot R, et al. Host preferences of Palaearctic *Culicoides* biting midges: implications for transmission of orbiviruses. *Med Vet Entomol.* 2013;27(3):255–66. <https://doi.org/10.1111/j.1365-2915.2012.01042.x> PMID: 22985009
- Elbers ARW, Meiswinkel R. *Culicoides* (Diptera: Ceratopogonidae) and livestock in the Netherlands: comparing host preference and attack rates on a Shetland pony, a dairy cow, and a sheep. *J Vector Ecol.* 2015;40(2):308–17. <https://doi.org/10.1111/jvec.12169> PMID: 26611966
- Downe AER. Blood-meal sources and notes on host preferences of some *Aedes* mosquitoes (Diptera: Culicidae). *Can J Zool.* 1960;38(4):689–99. <https://doi.org/10.1139/z60-072>
- Oyewole IO, Awolola TS. Impact of urbanisation on bionomics and distribution of malaria vectors in Lagos, southwestern Nigeria. *J Vector Borne Dis.* 2006;43(4):173.
- Tolulope O. Spatio-temporal clustering of malaria morbidity in Nigeria (2004–2008). *J Sci Res.* 2014;13(1):99–113.
- Dye C, Hasibeder G. Population dynamics of mosquito-borne disease: effects of flies which bite some people more frequently than others. *Trans R Soc Trop Med Hyg.* 1986;80(1):69–77. [https://doi.org/10.1016/0035-9203\(86\)90199-9](https://doi.org/10.1016/0035-9203(86)90199-9) PMID: 3727001
- Ross A, Smith T. Interpreting malaria age-prevalence and incidence curves: a simulation study of the effects of different types of heterogeneity. *Malar J.* 2010;9:132. <https://doi.org/10.1186/1475-2875-9-132> PMID: 20478060

25. Cooper L, Kang SY, Bisanzio D, Maxwell K, Rodriguez-Barraquer I, Greenhouse B, et al. Pareto rules for malaria super-spreaders and super-spreading. *Nat Commun.* 2019;10(1):3939. <https://doi.org/10.1038/s41467-019-11861-y>
26. Selvaraj P, Wenger EA, Gerardin J. Seasonality and heterogeneity of malaria transmission determine success of interventions in high-endemic settings: a modeling study. *BMC Infect Dis.* 2018;18(1):413. <https://doi.org/10.1186/s12879-018-3319-y> PMID: 30134861
27. Smith T, Maire N, Dietz K, Killeen GF, Vounatsou P, Molineaux L, et al. Relationship between the entomologic inoculation rate and the force of infection for *Plasmodium falciparum* malaria. *Am J Trop Med Hyg.* 2006;75(2 Suppl):11–8. https://doi.org/10.4269/ajtmh.2006.75.2_suppl.0750011 PMID: 16931811
28. Guelbéogo WM, Gonçalves BP, Grignard L, Bradley J, Serme SS, Hellewell J, et al. Variation in natural exposure to anopheles mosquitoes and its effects on malaria transmission. *Elife.* 2018;7:e32625. <https://doi.org/10.7554/eLife.32625> PMID: 29357976
29. Gonçalves BP, Kapulu MC, Sawa P, Guelbéogo WM, Tiono AB, Grignard L, et al. Examining the human infectious reservoir for *Plasmodium falciparum* malaria in areas of differing transmission intensity. *Nat Commun.* 2017;8(1):1133. <https://doi.org/10.1038/s41467-017-01270-4> PMID: 29074880
30. Stan. Stan user's guide version 2.36. 2024. Available from: https://mc-stan.org/docs/2_36/stan-users-guide/index.html
31. RStudio Team. Rstudio: Integrated Development Environment for R. Boston, MA: RStudio, PBC; 2020. <http://www.rstudio.com/>
32. Gronau QF, Singmann H, Forster JJ, Wagenmakers EJ, Team TJ, Guo J, et al. Bridgesampling: Bridge Sampling for Marginal Likelihoods and Bayes Factors. 2021. <https://github.com/quentingronau/bridgesampling>
33. Chitnis N, Hardy D, Smith T. A periodically-forced mathematical model for the seasonal dynamics of malaria in mosquitoes. *Bull Math Biol.* 2012;74(5):1098–124. <https://doi.org/10.1007/s11538-011-9710-0> PMID: 22218880
34. Chitnis N, Smith T, Steketee R. A mathematical model for the dynamics of malaria in mosquitoes feeding on a heterogeneous host population. *J Biol Dyn.* 2008;2(3):259–85. <https://doi.org/10.1080/17513750701769857> PMID: 22876869
35. Smith T, Maire N, Ross A, Penny M, Chitnis N, Schapira A, et al. Towards a comprehensive simulation model of malaria epidemiology and control. *Parasitology.* 2008;135(13):1507–16. <https://doi.org/10.1017/S0031182008000371> PMID: 18694530
36. Reiker T, Golumbeanu M, Shattock A, Burgert L, Smith TA, Filippi S, et al. Emulator-based Bayesian optimization for efficient multi-objective calibration of an individual-based model of malaria. *Nat Commun.* 2021;12(1):7212. <https://doi.org/10.1038/s41467-021-27486-z> PMID: 34893600
37. Galactionova K, Tediosi F, de Savigny D, Smith T, Tanner M. Effective coverage and systems effectiveness for malaria case management in sub-Saharan African countries. *PLoS One.* 2015;10(5):e0127818. <https://doi.org/10.1371/journal.pone.0127818> PMID: 26000856
38. Battle KE, Guerra CA, Golding N, Duda KA, Cameron E, Howes RE, et al. Global database of matched *Plasmodium falciparum* and *P. vivax* incidence and prevalence records from 1985–2013. *Sci Data.* 2015;2:150012. <https://doi.org/10.1038/sdata.2015.12>
39. Kamber L, Cavelan A, Rourre A, Chitnis N, Penny M. Application of multi-objective Bayesian optimization methods to complex individual-based models: An updated calibration of OpenMalaria. 2025.
40. Smith DL, Dushoff J, McKenzie FE. The risk of a mosquito-borne infection in a heterogeneous environment. *PLoS Biol.* 2004;2(11):e368. <https://doi.org/10.1371/journal.pbio.0020368> PMID: 15510228
41. Hasibeder G, Dye C. Population dynamics of mosquito-borne disease: persistence in a completely heterogeneous environment. *Theor Popul Biol.* 1988;33(1):31–53. [https://doi.org/10.1016/0040-5809\(88\)90003-2](https://doi.org/10.1016/0040-5809(88)90003-2) PMID: 2897726